

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-100801

(43)Date of publication of application : 23.04.1993

(51)Int.Cl.

G06F 3/06

G06F 3/06

(21)Application number : 03-261820

(71)Applicant : HITACHI LTD

(22)Date of filing : 09.10.1991

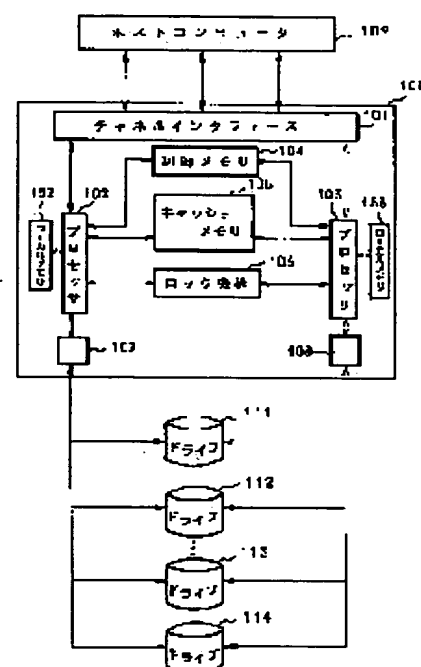
(72)Inventor : MURATA TOMOHIRO
KURIHARA KENZO
AKATSU MASA HARU
SHIRAYANAGI YOSHIRO

(54) SYSTEM FOR MAKING DISK SUB-SYSTEM HIGHLY RELIABLE

(57)Abstract:

PURPOSE: To obtain a method for improving fault resistance against the access fault of a drive in a large disk sub-system connecting multiple drives.

CONSTITUTION: When statistical information on the occurrence of an access error which can be corrected on a present-use drive (track) is collected and it is judged that there is the omen of the fault of the present-use drive (track) based on said information, a disk controller 100 copies data from the present-use drive (track) to a standby drive (track) and switches the present-use drive (track) to the standby drive (track). Since the disk controller 100 saves data on the present-use drive (track) which is judged that there is the omen of the fault to the standby drive (track) and uses the standby drive (track) as the new present-use drive (track), the fault of the present-use drive can be avoided preventively. Thus, the possibility of the drive fault reduces and the reliability of the disk sub-system improves.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection][Date of requesting appeal against examiner's decision
of rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平5-100801

(43)公開日 平成5年(1993)4月23日

(51)Int.Cl.⁵

G 0 6 F 3/06

識別記号

3 0 6 F 7165-5B

3 0 4 F 7165-5B

庁内整理番号

7165-5B

7165-5B

F I

技術表示箇所

審査請求 未請求 請求項の数 2 (全 17 頁)

(21)出願番号

特願平3-261820

(22)出願日

平成3年(1991)10月9日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 村田 智洋

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

(72)発明者 栗原 謙三

神奈川県川崎市麻生区王禅寺1099番地 株
 式会社日立製作所システム開発研究所内

(72)発明者 赤津 雅晴

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

(74)代理人 弁理士 小川 勝男

[最終頁に続く](#)

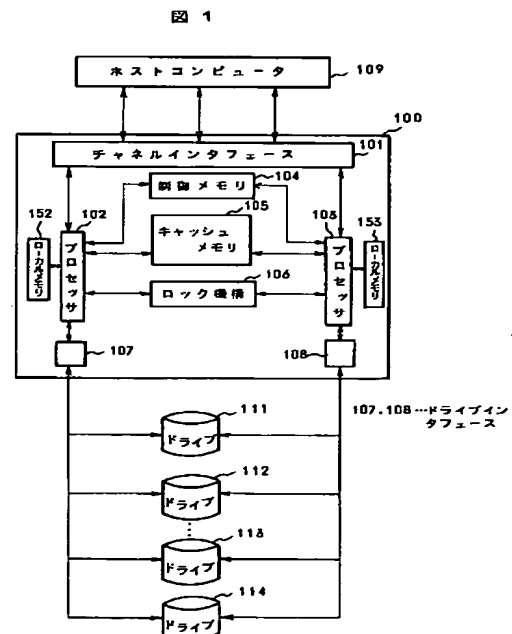
(54)【発明の名称】 デイスクサブシステムの高信頼化方式

(57) 【要約】

【目的】本発明はディスク制御装置に関し、特に、多数のドライブを接続する大型ディスクサブシステムにおいて、ドライブのアクセス障害に対する耐障害性を向上させる方法に関する。

【構成】現用ドライブ（トラック）に関する訂正可能なアクセスエラー発生に関する統計情報を収集し、それをもとに、現用ドライブ（トラック）障害の前兆があると判断した場合に、ディスク制御装置が現用ドライブ（トラック）から予備ドライブ（トラック）へのデータをコピー後、現用ドライブ（トラック）から予備ドライブ（トラック）への切り替えを行なう事により達成される。

【効果】ディスク制御装置が、障害の前兆があると判断された現用ドライブ（トラック）のデータを、予備ドライブ（トラック）に退避した後、予備ドライブ（トラック）を新たな現用ドライブ（トラック）として使用するため、現用ドライブの障害を予防的に回避することができ、ドライブ障害の確率が減少し、ディスクサブシステムの信頼性が向上する。



【特許請求の範囲】

【請求項１】 ディスク制御装置と、磁気記憶媒体であるドライブからなるディスクサブシステムにおいて、ディスク制御装置に接続するドライブを、一台以上の現用ドライブと一台以上の予備ドライブに分けてディスク制御装置にあらかじめ登録しておき、現用ドライブに対するホストコンピュータからの入出力実行時に発生した訂正可能なデータ転送系障害の回数をディスク制御装置でドライブごとに計数し、該発生回数があらかじめ定められた値を越えた現用ドライブについて、現用ドライブの内容をディスク制御装置が適当な予備ドライブにコピーし、該予備ドライブを該現用ドライブの代替として使用することを特徴とするディスクサブシステムの高信頼化方式。

【請求項２】 請求項１記載のディスク制御装置において、ドライブが有するトラックを現用トラックと、交替トラックに分けてディスク制御装置にあらかじめ登録しておき、現用トラックに対するホストコンピュータからの入出力実行時に発生した訂正可能なデータ転送系障害の回数をディスク制御装置でトラックごとに計数しておき、該発生回数があらかじめ定められた値を越えた現用トラックについて、ディスク制御装置が同一ドライブ内の適当な予備トラックにその内容をコピーした後、該現用トラックの代替として該交替トラックを現用トラックとして使用することを特徴とするディスクサブシステムの高信頼化方式。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】 本発明はディスク制御装置に関し、特に、多数のドライブを接続する大型ディスクサブシステムにおいて、ドライブのアクセス障害に対する耐障害性を向上させる方法に関する。

【０００２】

【従来の技術】 ホストコンピュータからのディスクドライブに対する一回のデータ書き込みで、ディスク制御装置が、正、副ペアドライブに同一データを書き込み、正、副いずれかのドライブでアクセス障害が発生しても、正常な方のドライブでアクセスを継続することにより、ドライブのアクセス障害に対する耐障害性を向上させる方法（ディスク制御装置による二重書き）がアクセス ３月／４月号（１９８８年）第９頁から第１０頁（ACCESS MAR/APR １９８８ pp 9-10）に記載されている。

【０００３】

【発明が解決しようとする課題】 正、副ペアドライブによる二重書きは、耐障害性が高いがコストも大きい。このため、オンライン処理系のドライブに適用が限定される場合が多い。すなわち、アクセス頻度はそう多くないが、一旦アクセス障害が発生するとその回復処理のための影響が大きいオフライン処理系のドライブについて

も、耐障害性向上の要求はあるものの、一般にその数が多いため、二重書きの適用はコスト面で難しい。

【０００４】 本発明の目的は、上記オフライン処理系のドライブを対象に、予防保守を行なうことにより、その障害の確率を減らす方法を提供することである。

【０００５】

【課題を解決するための手段】 上記の目的は、現用ドライブ（トラック）に関する訂正可能なアクセスエラー発生に関する統計情報を収集しておき、それをもとに、現用ドライブ（トラック）障害の前兆があると判断した場合には、ディスク制御装置が現用ドライブ（トラック）から予備ドライブ（トラック）へのデータコピー後、現用ドライブ（トラック）から予備ドライブ（トラック）への切り替えを行なう事により達成される。

【０００６】

【作用】 障害の前兆があると判断された現用ドライブ（トラック）のデータを、ディスク制御装置が予備ドライブ（トラック）にコピーした後、予備ドライブ（トラック）を新たな現用ドライブ（トラック）として使用することで、現用ドライブの障害を予防的に回避する。

【０００７】

【実施例】 図１に、本発明をディスク制御サブシステムに適用した場合の一実施例を示す。図１において、ディスク制御装置１００は、上位側でホストコンピュータ１０９と接続され、また下位側で記憶媒体であるドライブ１１１、１１２、１１３、１１４と接続される。ディスク制御装置１００は、これらのドライブ上でホストコンピュータ１０９の要求に応じてデータのリード、ライトを行なう。以下、図１におけるディスク制御装置１００の構成を参照しながら、本発明の主内容であるディスク制御装置１００の動作を説明する。

【０００８】 各ドライブとホストコンピュータ１０９との間のデータ転送を実際に行なうのはディスク制御装置１００に内蔵されたプロセッサ１０２、１０３である。これらのプロセッサは、チャネルインタフェース１０１、ドライブインタフェース１０７、１０８を介し、ホストコンピュータ１０９、及び各ドライブとそれぞれ接続する。制御メモリ１０４はすべてのプロセッサからアクセス可能な共通メモリであり、ディスク制御装置１００がドライブをアクセスするための共通制御情報が格納される。共通制御情報の内容については、以下で必要に応じて説明する。

【０００９】 図２は共通制御情報であるドライブ制御ブロック２００（Device Control Block；略してDCB）を示す。DCBには、各デバイスをディスク制御装置１００が識別するためのドライブ番号２０１（Device Connection Address；略してDCA）、当該ドライブの型式の種類を示すドライブタイプ各ドライブを２０２が格納される。また、ホストコンピュータの要求に応じてディスク制御装置１００が実際にデータの入出力を行なう現

用ドライブと、現用ドライブのバックアップ用である予備ドライブの2種類に各ドライブを区別し、当該ドライブがそのどちらであるかを用途区別（現用／予備）203に格納する。その他のDCB制御情報の内容については後述する。

【0010】図3は共通制御情報である交替トラック管理情報300を示す。各ドライブごとに、データを格納する現用トラックのバックアップとして使用可能な未使用予備トラックと、すでに現用トラックのバックアップとして使用中である予備トラックを区分して管理する。具体的には、予備トラックのトラック番号を格納した交替トラック制御ブロックを、未使用と使用中に分けてキューで管理する。

【0011】図4は共通制御情報であるトラックアクセス・テンポラリエラー統計情報800である。ディスク制御装置100内の各プロセッサは、各ドライブのトラック上のデータリード処理において、訂正可能データ転送エラーを検出する度に、トラックアクセス・テンポラリエラー統計情報800の該当ドライブ801の該当するトラック欄802に格納されているテンポラリー・アクセスエラー合計回数803をインクリメントする。これらの情報は、トラック又はドライブの障害発生の際を検出するために用いられるが、詳細については後述する。

【0012】続いて、本発明におけるディスク制御装置100内の各プロセッサの動作を説明する。図5はドライブヘルスチェック処理のフローを示す。本処理はディスク制御装置100内の各プロセッサにより、ある周期で繰返し実行される。最初に共通制御情報であるチェックドライブ番号格納テーブル800（図8）にドライブ番号として0を設定しておく。まず、チェックドライブ番号格納テーブル800より、チェックすべきドライブの番号を読みだし、自プロセッサのローカルメモリ（152、153）に格納する（501）。

【0013】次に、当該ドライブ番号を有するドライブのスイッチ要求テーブル900（図9）のスイッチ要求901が‘off’であることをチェックする（502）。ここで、スイッチ要求テーブル900において、スイッチ要求901が‘off’であるとは、当該ドライブに関して予備ドライブ又は、予備トラックへの切り替え要求がないことを示す。スイッチ要求901が‘off’である場合は当該ドライブ番号が最大ドライブ番号に等しいか否かをチェックし、等しい場合はチェックドライブ番号格納テーブル800に-1を設定しておく。

【0014】続いて、ローカル変数であるトラックサーチポインタをゼロクリアし（505）、先に自プロセッサのローカルメモリに格納しておいたチェックドライブ番号を持つドライブについて、上記トラックサーチポインタと等しいトラック番号を有するトラックのテンポラリーアクセスエラー回数が上限値（スレシールド）を越えているか否かを判定し、越えている場合は当該ドライブの

テンポラリエラー多発トラック数合計804を+1し（514）、当該ドライブのテンポラリエラー多発トラック数合計804が上限値（スレシールド）を越えているか否かを判定し、越えている場合は当該ドライブの処理済みマーク801をセットした後、当該ドライブのスイッチ要求を登録し（511）、チェックドライブ番号を+1して処理を終わる。もし、上限値を越えていなければ上記トラックの処理済みマーク805をセットした後、当該トラックのスイッチ要求を登録し（509）、トラックサーチポインタを+1して次のトラックのヘルスチェックに行く。

【0015】図6は図5におけるトラックスイッチ要求登録処理509の処理内容を示すフローである。まず、スイッチ処理の対象となるトラックを持つドライブのロックを取得し（601）、未使用の交替トラックキュー310（図3）からコピー先となる交替トラックを一本選ぶ（604）。もし、未使用の交替トラックがなければトラックスイッチ不可を表示し（606）、取得したロックを解放して処理を終わる。未使用の交替トラックがあれば、スイッチ対象ドライブのDCBに交替トラック番号204を、またコピー先ドライブ番号205として自ドライブ番号を設定する。さらにコピー要求206をトラックコピーに設定し、コピー開始トラック212にスイッチ対象トラック番号を設定する。そしてスイッチ要求208をonし（708）、取得していたロックを解放する。

【0016】図7は図5におけるドライブスイッチ要求登録処理511の処理内容を示すフローである。まず、スイッチ処理の対象となるドライブのロックを取得し（701）、ディスク制御装置100に接続するドライブのDCBを調べ、用途区別203が予備ドライブであり、かつスイッチ対象ドライブとドライブタイプが一致するドライブを一つ選ぶ（704）。予備ドライブがなければスイッチ不可を表示し（706）、上記ロックを解除して処理を終わる。予備ドライブがあれば、スイッチ対象ドライブのDCBのコピー先ドライブ番号205に上記予備ドライブ番号を設定し、さらにコピー要求206をドライブコピーに設定し、コピー開始トラック212、及びコピー終了トラック213に上記ドライブタイプに対応したコピー開始トラック番号とコピー終了トラック番号をそれぞれ設定する（707）。そしてスイッチ要求208をonし（708）、取得していたロックを解放する。

【0017】図10はトラックスイッチ処理の処理内容を示すフローである。本処理も任意のプロセッサで、ある周期ごとに実行される。まず、スイッチ処理ドライブ番号格納テーブル900よりスイッチ処理の候補となるドライブ番号を読みだし、プロセッサのローカルメモリに格納する（1001）。当該ドライブ番号が最大ドライブ番号に等しければ当該テーブル900をゼロクリア

し(1003)、等しくなければ当該テーブル900を+1する(1015)。続いて、当該ドライブのDCB内のスワップ要求208が‘on’か否かをチェックし、‘on’でなければ処理を終わる。‘on’であれば、当該ドライブのコピー要求がトラックコピーかドライブコピーかを判定し、ドライブコピーであればドライブスワップ処理を行なって処理を終わる。トラックコピーであれば、スワップ要求208を処理中にし(1006)、当該ドライブのロックを取得し(1007)、コピー開始トラック番号を有するトラックからトラックデータをキャッシュメモリ105にロードし(1008)、交替トラック番号を有するトラックに対し、上記ロードデータをストアした後(1010)、スワップ処理要求を‘off’に設定し(1010)取得していた上記ロックを解放して処理を終わる。

【0018】図11はドライブスワップ処理の処理内容を示すフローである。本処理も任意のプロセッサで、ある周期ごとに実行される。まず、スワップ要求208を処理中にし、コピー済トラック番号214を(コピー開始番号-1)の値に初期化する(1101)。コピー済トラック番号214がコピー終了トラック番号に等しくなるまで以下を繰り返す。まず、コピー元ドライブ(当該ドライブ)、およびコピー先ドライブのロックをともに取得し(1103)、未コピートラック(コピー済トラック+1の番号を有するトラック)からトラックデータをキャッシュメモリ105にロードし(1105)、コピー先ドライブのトラックに上記ロードデータをストアした後(1010)、取得したロックを解放しコピー済トラックの番号を+1する(1107)。そして、ホストコンピュータからの、スワップ処理中の当該ドライブに対するデータ入出力を受け付け可能とするために、一定時間スワップ処理を中断した後(1108)、次のトラックのコピー処理に移る。

【0019】コピー済トラック番号214がコピー終了トラック番号に等しくなった時点で、コピー元、およびコピー先ドライブのロックを共に取得し、コピー先ドライブへの差分データのデステージ処理(1111)、及び、アクセスパススワップ処理(1112)を行ない、最後に、取得していたロックを解放し(1114)処理を終わる。

【0020】以下、差分データのデステージ処理(1111)の内容を説明するが、その前に、ドライブスワップ処理中の差分データの管理について述べておく。ホストコンピュータから、コピー元ドライブのあるトラックに書き込み要求が来た場合、当該トラックの番号がコピー済トラック番号より大きいか等しい場合は、データ書き込みの対象となるトラックは未コピーのトラックであり、当該トラックに単にデータの書き込みを行なえばよく、コピー終了時点で、コピー元とコピー先のドライブでデータが異なることはない。一方、当該トラックの番

号がコピー済トラック番号より小さい場合は、データ書き込みの対象となるトラックはコピー済のトラックであるため、当該トラックにデータの書き込みを行なうと、コピー終了時点でコピー元とコピー先のドライブでデータが異なる。

【0021】後者の場合に対処するために、図14に示すように、コピー元ドライブ上のコピー済のトラックにデータの書き込みを行なった場合には、当該トラック(このようなトラックを差分トラックと呼ぶ)に対する書き込みデータをキャッシュ105内のキャッシュ本体402にも保持する。例えば、トラック2(TR2)に対する書き込みがこのケースとする。この場合、まず、キャッシュ本体402内の空いているデータ格納場所のアドレスを保持しているスロット制御ブロック(SCB)のキュー43に有ったSCB(406)を一つ選び、そのSCBが指す格納場所(スロット2)に上記書き込みデータを格納し、キャッシュ105内のディレクトリ401に設けたトラック表403の、TR2に対応するエントリに、上記SCBへのポインタを格納する。このようにすることで、差分トラックがどれで、その書き込みデータが何であるかが、キャッシュを見れば分かる。差分トラックのデータは、ドライブコピーが一通り終了した段階、すなわち、コピー済トラック番号214がコピー終了トラック番号に等しくなった時点で、再度コピー先ドライブに書き込みを行なう必要がある。

【0022】さて、図12は、差分データのデステージ処理(1111)の内容を示すフローである。まず、上記ディレクトリ401のトラック表403に、上記差分トラックがリンクしているか否かチェックし(1201)、もし、一本もリンクしていなければ処理を終わる。リンクしている場合は、差分トラックのSCBが指すキャッシュ上の書き込みデータをすべて、コピー先ドライブの該当するトラックに書き込む(1204)。

【0023】次に、アクセスバスマッピングテーブル1400(図15)について説明する。アクセスバスマッピングテーブル1400(図15)は、ホストコンピュータ109から見たドライブの識別番号CCA(Channel Connection Address)と、ディスク制御装置100で内部的に管理するドライブの識別番号DCA(Device Connection Address)の対応表を与えるものである。ホストコンピュータ109の入出力要求で指示されるドライブ番号(CCA)は、ディスク制御装置100によりDCAに変換され、DCAに対応するドライブが実際にアクセスされる。具体的には図15で、CCAの値をkとすると、アクセスバスマッピングテーブル1400のCCAマッピング情報1401のBkの欄の値iを読みだしてDCAの値とし、ドライブ番号iのドライブが実際にアクセスされる。このマッピング機能により、上記のドライブスワップ処理で現用ドライブが予備ドライブに切り替えられても、CCAとDCAの対応関係を変えるこ

とにより、ホストコンピュータ側でドライブ番号を変えずにすむ。

【0024】最後に、上記のアクセスバスマッピングテーブル1400を用いたアクセスバスマッピング処理について、図13のフローに従い説明する。まず、コピー元ドライブのDCA（iとする）とコピー先ドライブのDCA（jとする）を記憶する（1301）。次に、DCAマッピング情報1402のB_iの欄の値（kとする）と、DCAマッピング情報1402のB_jの欄の値（lとする）をリードする（1302）。続いて、DCAマッピング情報1402のB_iの欄に値lを書き込み、CCAマッピング情報1401のB_lの欄に値iを書き込む（1303）。最後に、DCAマッピング情報1402のB_jの欄に値kを書き込み、CCAマッピング情報1401のB_kの欄に値jを書き込む（1303）。

【0025】これらの処理により、障害の前兆があると判断された現用ドライブ（トラック）のデータを、ディスク制御装置が予備ドライブ（トラック）にコピーした後、予備ドライブ（トラック）を新たな現用ドライブ（トラック）として使用することが自動的に実現でき、ディスクサブシステムの信頼性が向上する。

【0026】

【発明の効果】障害の前兆があると判断された現用ドライブ（トラック）のデータを、ディスク制御装置が予備ドライブ（トラック）にコピーした後、予備ドライブ（トラック）を新たな現用ドライブ（トラック）として使用するため、現用ドライブの障害を予防的に回避する

ことができ、ドライブ障害の発生確率が減少し、ディスクサブシステムの信頼性が向上する。

【図面の簡単な説明】

【図1】本発明の一実施例を示す制御構成図である。

【図2】ドライブ制御ブロックの構成図である。

【図3】交替トラックの管理情報である。

【図4】トラックアクセス・テンポラリエラー統計情報を示す図である。

【図5】ドライブヘルスチェック処理のフローである。

【図6】トラックスワップ要求登録処理のフローである。

【図7】ドライブスワップ要求登録処理のフローである。

【図8】チェックドライブ番号格納テーブル。

【図9】スワップ処理ドライブ番号格納テーブル。

【図10】トラックスワップ処理のフローである。

【図11】ドライブスワップ処理のフローである。

【図12】差分トラックデステージ処理のフローである。

【図13】アクセスバスマッピング処理のフローである。

【図14】キャッシュメモリの構成図である。

【図15】アクセスバスマッピングテーブルの構成図である。

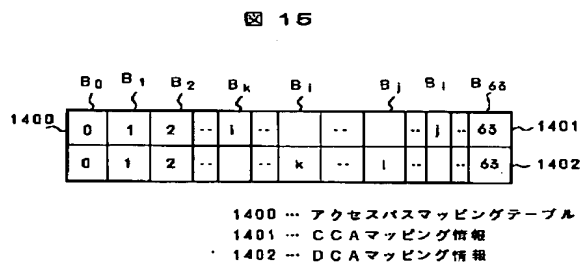
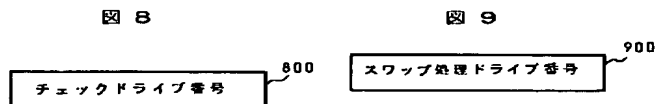
【符号の説明】

100…ディスク制御装置、102、103…プロセッサ、104…制御メモリ、109…ホストコンピュータ、111、112、113、114…ドライブ。

【図8】

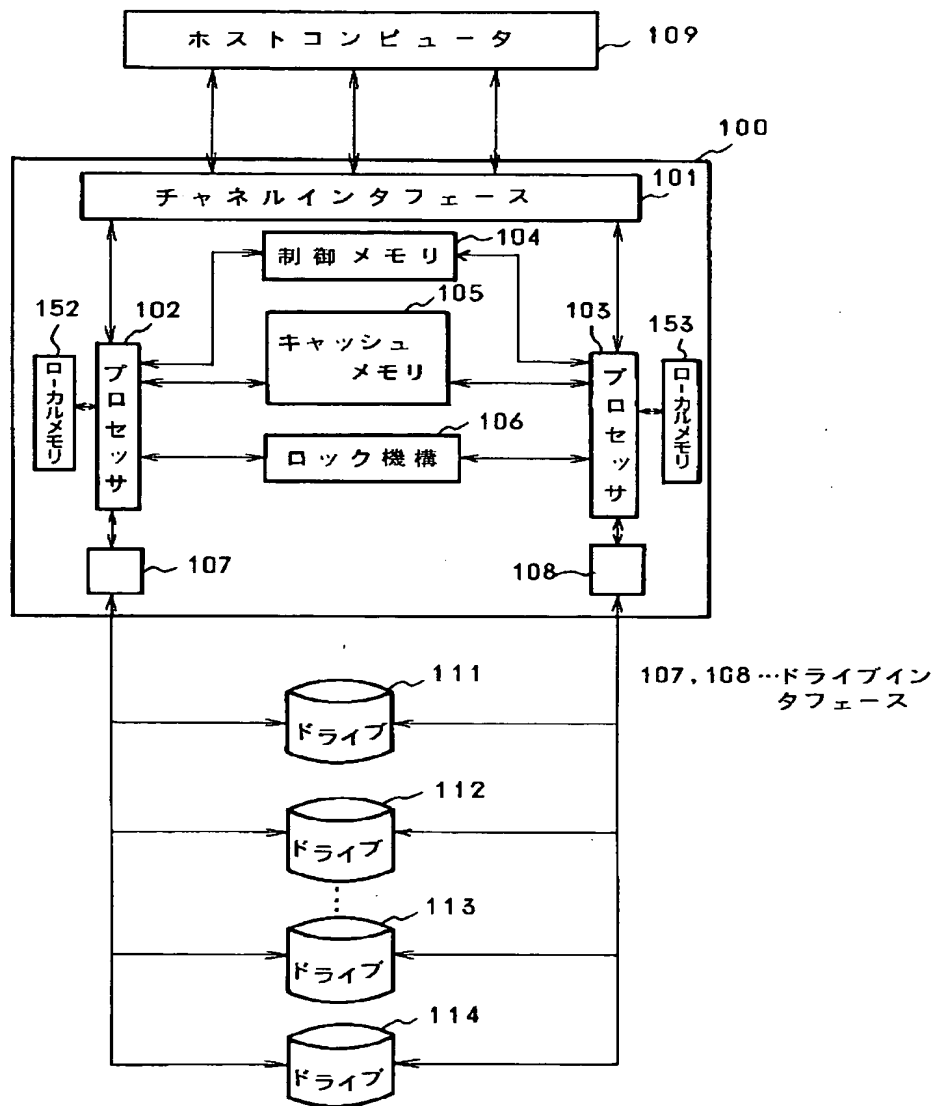
【図9】

【図15】



【図1】

図 1



【図2】

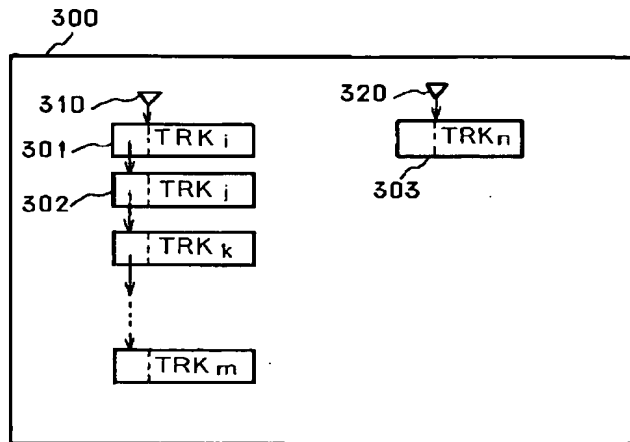
図 2

200	
ドライブ番号 (D C A)	201
ド ラ イ ブ タ イ プ	202
用途区別 (現用 / 予備)	203
交替用予備トラック番号	204
コピー先ドライブ番号	205
コピー要求 (トラック / ドライブ)	206
デ ス テ ー ジ 要 求	207
スワップ要求 (有 / 無 / 処理中)	208
ド ラ イ ブ 保 守 要 求	209
チャンネルへの 割込要求	210
テンポラリアクセスエラー多発トラック 合計数	211
コピー開始トラック番号	212
コピー終了トラック番号	213
コピー済トラック番号	214

300... ドライブ制御ブロック (D C B)

【図3】

図 3



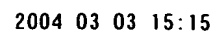
- 300 交替トラック管理情報
- 301, 302, 303 ... 交替トラック制御ブロック
- 310 未使用交替トラックキュー
- 320 使用中交替トラックキュー

【図4】

図 4

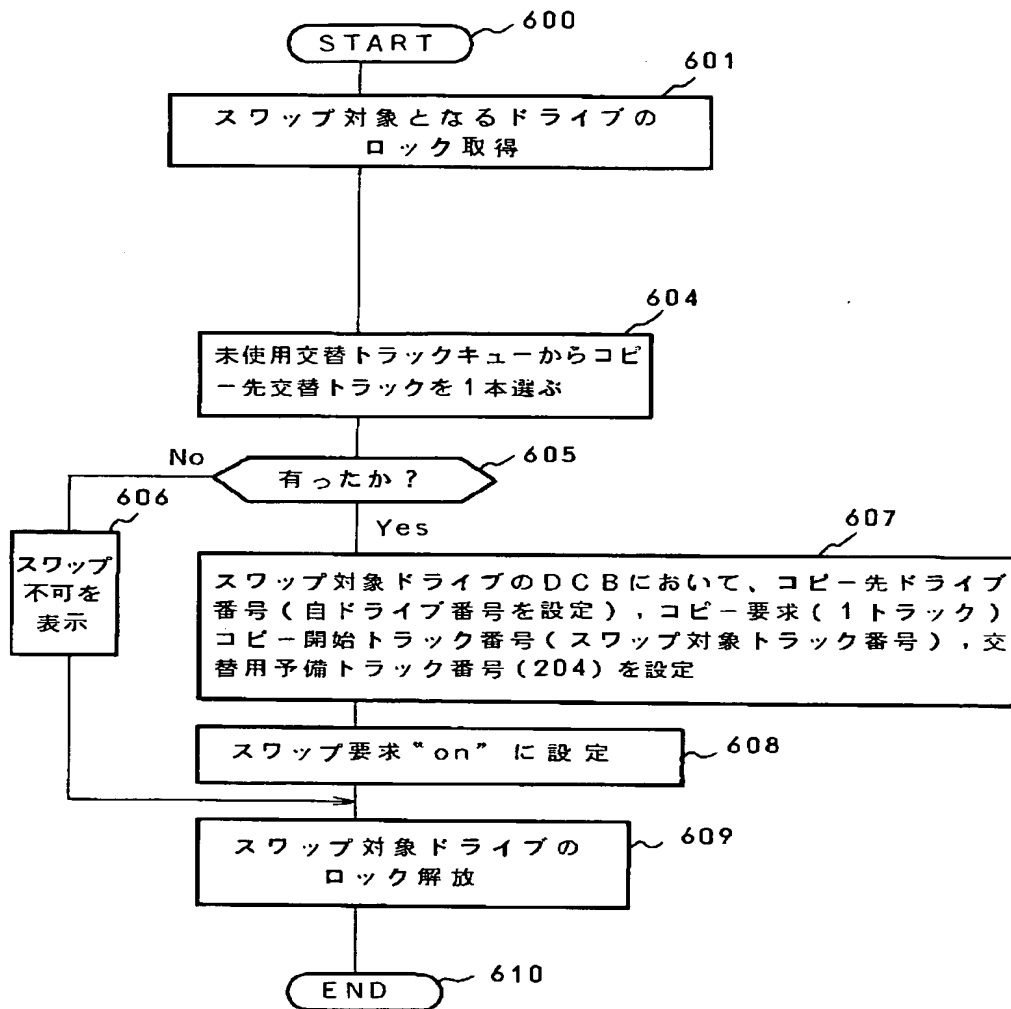
ドライブ 番号	処理 済	トラック 番号	処理 済	テンポラリアクセスエラー 合計回数	エラー多発 トラック数合計
0		0		1 00	S ₀
		1		1 01	
		2		1 02	
		⋮		⋮	
		m		1 0m	
1		0		1 10	S ₁
		1		1 11	
		⋮			
		n		1 1m	
⋮	⋮	⋮			
M		0		1 M0	S _M
		⋮			
		m		1 Mm	

图 5



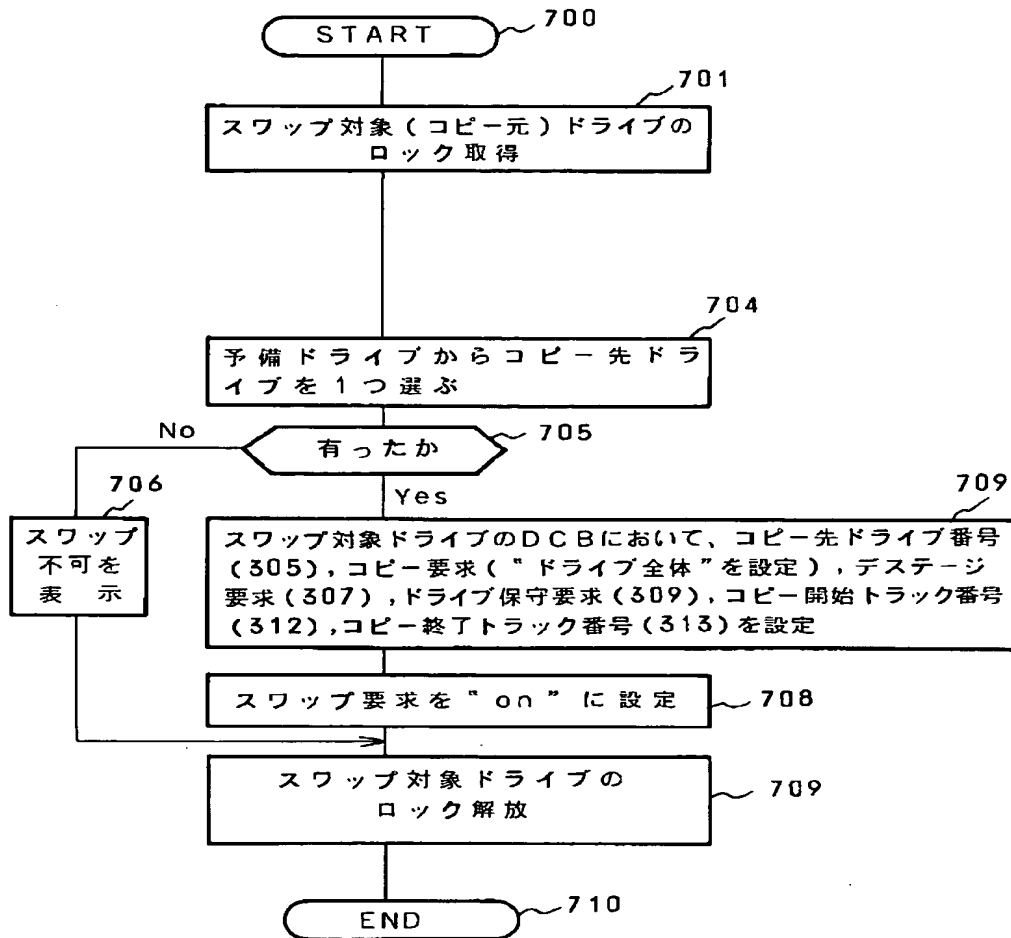
【図6】

図 6



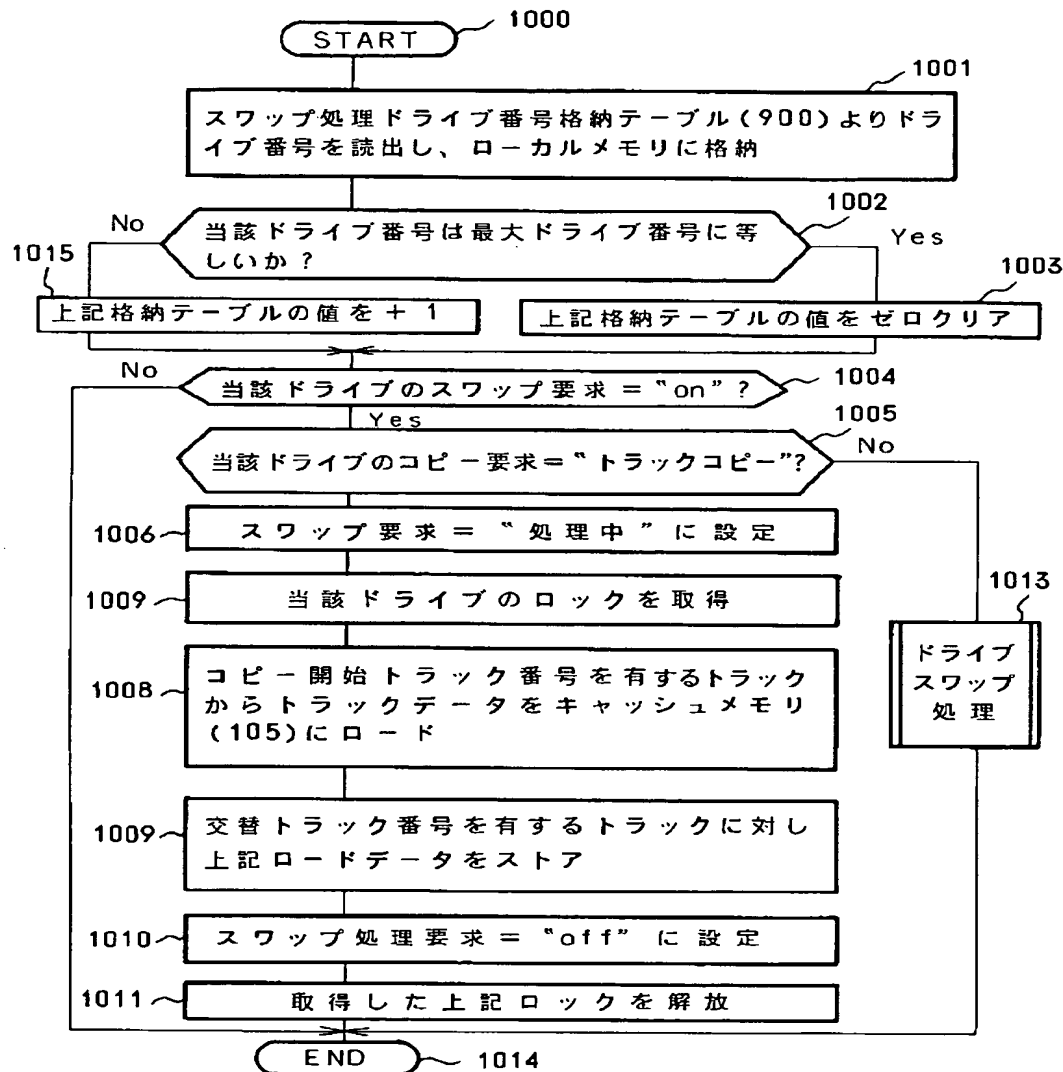
【図7】

図 7



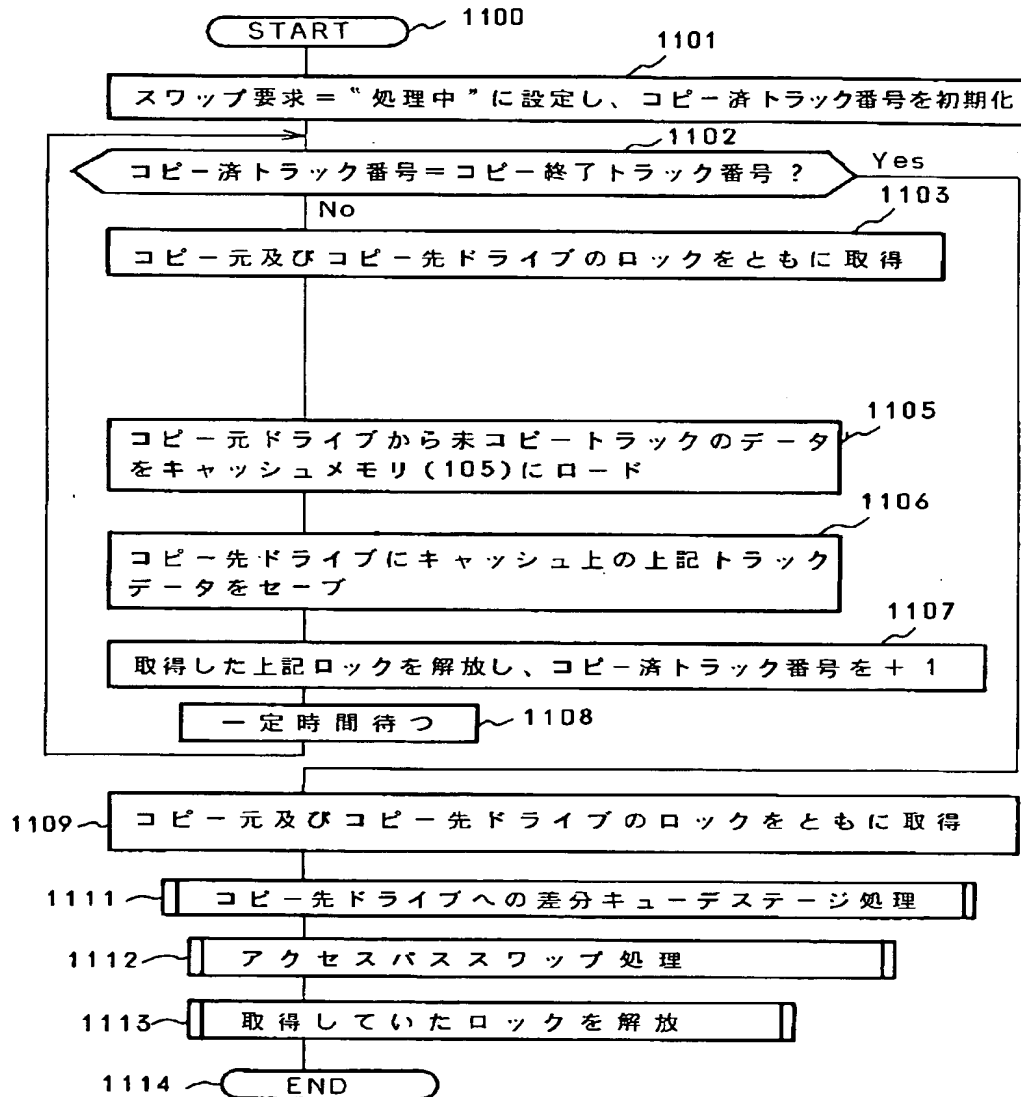
【図10】

図 10



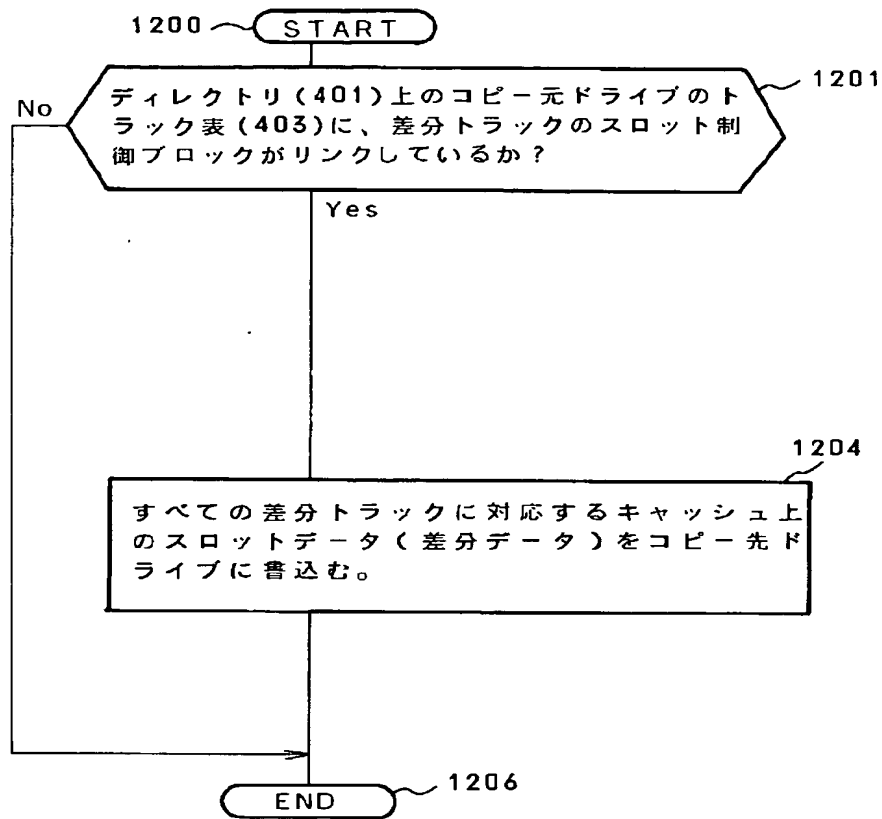
【図11】

図 11



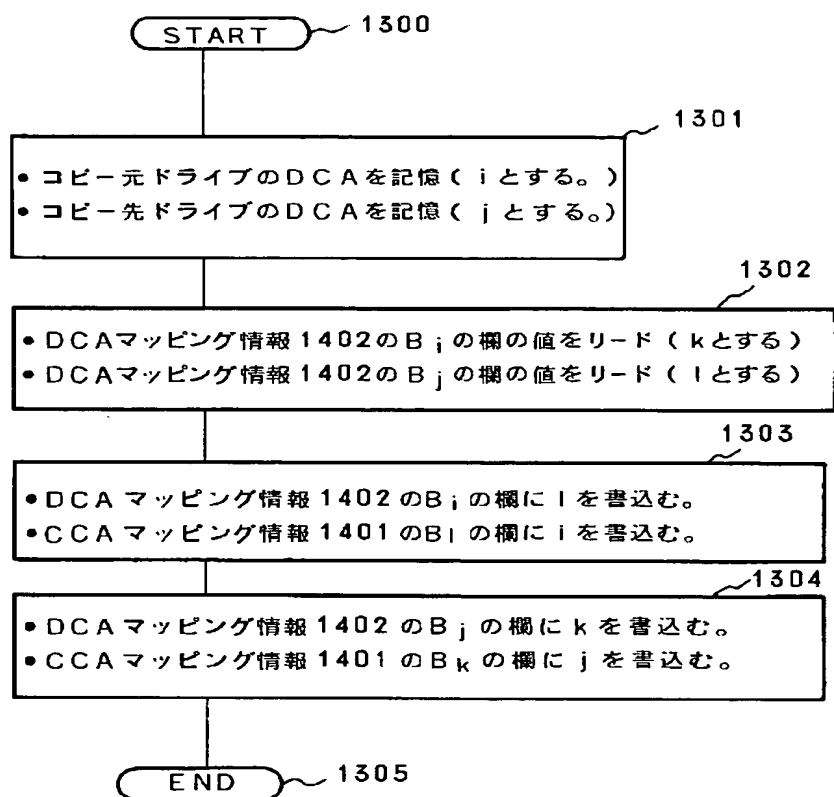
【図12】

図 12



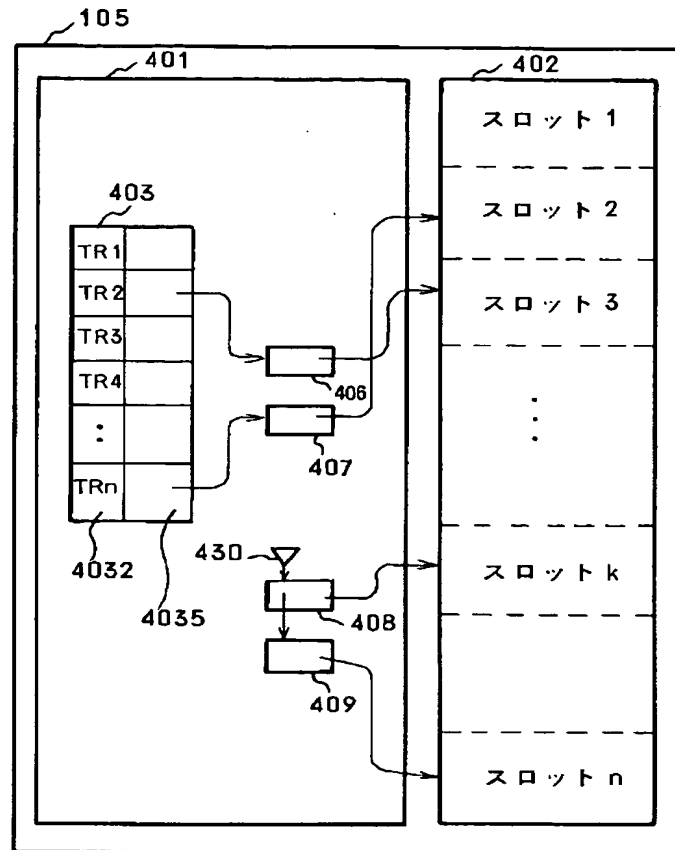
【図13】

図 13



【図14】

図 14



- 400 キャッシュメモリ
- 401 ディレクトリ
- 402 キャッシュ本体
- 403 トラック表
- 406 ~ 409 ... スロット制御ブロック
- 420 差分キュー
- 430 未使用キュー
- 4032 トラック番号
- 4034 交替トラックポインタ
- 4035 トラック表エントリ

フロントページの続き

(72)発明者 白柳 芳朗

45

神奈川県小田原市国府津2880番地 株式会
社日立製作所小田原工場内